

# Conditions for an experimental Boson-sampling computer to disprove the Extended Church-Turing thesis

V. S. Shchesnovich

*Centro de Ciências Naturais e Humanas, Universidade Federal do ABC, Santo André, SP, 09210-170 Brazil*

We consider an experimentally verifiable version of the Extended Church-Turing thesis and derive sufficient conditions for an imperfect experimental realization of the Boson-sampling computer of Aaronson & Arkhipov to disprove it. Since the effect of the photon mode mismatch and the network imperfections (noise) are already known, we focus on the multi-photon and vacuum components of the input, the photons losses, and the detector dark counts. The main result is the complete set of sufficient conditions on the experimental imperfections for scalability of the experimental Boson-sampling computer.

PACS numbers: 03.67.Lx, 05.30.Jp, 42.50.Ar

*Introduction* – The boson-sampling (BS) computer was recently proposed by S. Aaronson and A. Arkhipov [1] as a near-future feasible device serving as an evidence against the Extended Church-Turing thesis (ECT) stating that any physical device can be efficiently simulated on a classical computer. Such a device uses a unitary linear network with identical single photons at the network input. It is not known if the BS computer can solve any  $NP$  decision problem such as factoring large integers [2, 3]. What S. Aaronson and A. Arkhipov have shown is that simulation of the BS computer output on a classical computer would require exponential resources [1]. In the ideal case (with the ideally indistinguishable single photons, a perfect network and the ideal detectors) the  $N$ -boson output amplitudes are given as the matrix permanents (see Ref. [4]) of complex  $N \times N$ -submatrices of the network matrix [5, 6] and, by the classic result of the computation complexity theory [7], require exponential in  $N$  computation time (see also Ref. [8]). The fastest known algorithm for computation of the matrix permanent, due to H. Ryser [9], requires  $O(N^2 2^N)$  operations. A polynomial classical algorithm for the matrix permanent is believed to be impossible on the basis that the contrary would imply that all problems in the complexity class  $\#P$ , superior to the class  $NP$ , are efficiently solvable on a classical computer, i.e. the whole polynomial hierarchy of the computational complexity would collapse [1].

An universal quantum computer (UQC) could, of course, simulate the BS device, but the scalability of the BS beyond the classical computational power is easier to achieve. Indeed, with few dozens of single photons it would outperform the current classical computers [1]. Moreover, for this goal only passive optical elements and indistinguishable single photons [10] producing the Hong-Ou-Mandel type interference [11] (see also Refs. [12, 13]) are needed. Hence, though it is not known if any practical computational problem can be solved on the BS computer, such a device undoubtedly would have an enormous conceptual impact on physics. Four independent groups have tested their prototypes of the device on small networks with few single photons [14–17]. Now the goal

is to scale up the BS computer to at least few dozens of single photons.

In contrast with the UQC [18, 19], no specific error correction protocol is known for the BS device (i.e., not based on the UQC simulation of the BS computer). However, in Ref. [1] it was shown that *an approximate simulation* of the ideal BS computer output is also classically hard, if two highly plausible, numerically tested, conjectures are true. This result is valid for a network with  $M$  modes such that  $M \gg N^2$ . Based on this result, the scalability of the experimental BS computer can be studied. Some necessary, though not sufficient, conditions for the BS operation beyond the power of the classical computation were presented in Refs. [20, 21]. Two precise results are also available [22, 23]. First, it was shown [22] that the BS device employing a noisy optical network with the fidelity of its optical elements  $\mathcal{F}_{el} = 1 - O(N^{-2})$  is still hard to simulate classically. Second, it was also shown [23] that the BS computer is scalable when the average single-photon fidelity between any pair of the photon sources satisfies  $\langle \mathcal{F}_{ph} \rangle = 1 - O(N^{-3/2})$ . In a more recent study [24] some arguments were given that the BS device cannot serve as an evidence against the ECT, because it cannot be scaled up under a constant operation error. In this respect it agrees with the previous results [22, 23] implying the same conclusion.

*Experimentally verifiable ECT* – The conceptual importance of the BS computer for physics lies in its obvious conflict with the ECT [1]. Since the ECT thesis is an asymptotic statement, whereas the experimental verification uses finite devices, the *asymptotic* ECT cannot be rigorously disproven by any experiment. In the case of the UQC there are theoretical arguments for scalability to an arbitrarily large device (and arbitrary long computational time) using the error correction protocols [18, 19]. The BS computer hardness is not based on any decision problem in the  $NP$  class, but on the  $\#P$  class of higher computational complexity. Hence, the verification of its operation is classically hard. However, the most important point of Aaronson & Arkhipov [1] is that the unavoidable errors in the BS device are *an integral part* of the computational problem which is proven (un-

der two highly plausible conjectures) to be hard for the classical computers. Precisely, Theorem 1.3 of Ref. [1] states that given a random  $M$ -mode network with  $N$  of its input modes connected to the ideal single photon sources (with  $M \gg N^2$ ), it is classically hard to simulate the output distribution of this *ideal* BS computer in  $\text{poly}(N, 1/\epsilon)$  time, where  $\epsilon$  is the variational distance error. This computational problem can serve as a basis of the experimentally verifiable evidence against the ECT by the BS computer, if one shows that it can be efficiently simulated by the *experimentally feasible* BS device for any problem size  $N$  and any nonzero  $\epsilon$ .

Below a complete set of sufficient conditions is given for the experimental BS device with  $N$  photon sources to simulate the output of the ideal BS computer to a variational distance error  $\epsilon$  for the fraction  $1-\delta$  of all networks, for any finite  $N$  and any  $\epsilon, \delta > 0$ . Therefore, our results prove the scalability of the experimental BS computer, where one has to use  $1/(1-\delta)$  different networks, on average to have a hard instance of the network [32].

The asymptotic limit, i.e.  $N \rightarrow \infty$  and  $\epsilon \rightarrow 0$ , in practical terms implies that an arbitrarily small error has to be achievable. Therefore, one can consider separately the effect of various types of error. The photon mode mismatch and the noise in the network were already analyzed previously [22, 23]. To complete the picture, we will consider the combined effect of the multi-photon components and vacuum in the input modes, the photon losses, and the dark counts of the detectors.

*The experimental BS computer model* – We consider a  $M$ -mode unitary linear network connected to  $N$  identical sources (i.e. being replicas of each other) which output imperfect single photons with the density matrix  $\rho = p_0|0\rangle\langle 0| + p_1\rho^{(1)} + p_2\rho^{(2)} + \dots$ , where the  $k$ -photon component  $\rho^{(k)}$  appears with the probability  $p_k$ . As above discussed, we completely neglect the mode mismatch in each component  $\rho^{(k)}$ , thus  $\rho^{(k)} = |k\rangle\langle k|$  (note that the cross-coherence terms containing  $|k\rangle\langle m|$ , with  $m \neq k$ , are invisible to the photon detectors).

To account for the photon losses we assume them to occur at the detection stage. We assume the  $M$  photon-number unresolving (bucket) detectors, connected to the network output modes, to be replicas of each other. The genuine quantum features are embodied in the quantum output probabilities of the network, whereas the photon losses and the dark counts of the detectors have direct analogy in the classical particle counting with losses and count errors. The photon losses can be accounted for by introduction of the loss probability  $r$ , whereas the dark counts can be described by the integral dark count rate  $\nu$  (see details in Refs. [25, 26]). The  $M$  detectors are thus replicas of the bucket detector with the no-click probability  $P_D(0|s) = e^{-\nu}r^s$ , for the  $s$ -photon input (i.e., the zero dark counts probability  $e^{-\nu}$  multiplied by the total loss probability  $r^s$ ), and the click probability  $P_D(1|s) = 1 - e^{-\nu}r^s$ .

Let the input and output modes of the network have the boson operators  $a_i$  and  $b_i$ ,  $i = 1, \dots, M$ . These are re-

lated by the network matrix  $U$ :  $a_i^\dagger = \sum_{l=1}^M U_{il}b_l^\dagger$ . The input state (density matrix), with the input modes  $1, \dots, N$  connected to the photon sources, reads

$$\rho^{(in)} = \rho_1 \otimes \dots \otimes \rho_N \otimes |0\rangle\langle 0| \otimes \dots \otimes |0\rangle\langle 0|, \quad (1)$$

$$\rho_i \equiv \sum_{k=0}^{\infty} \frac{p_k}{k!} (a_i^\dagger)^k |0\rangle\langle 0| a_i^k.$$

It is convenient to introduce the vector notations for the mode occupation numbers, writing  $|\vec{n}\rangle$  for the Fock state with  $\vec{n} = (n_1, \dots, n_M)$ . Let us set  $|\vec{n}| = \sum_{i=1}^M n_i$ . In Eq. (1) we have the input  $\vec{n}$  with  $n_i \geq 0$ , for  $1 \leq i \leq N$ , and  $n_i = 0$ , for  $i \geq N+1$ . The input Fock state  $|\vec{n}, in\rangle$  can be expanded in the output Fock states  $|\vec{s}, out\rangle$  as follows [1, 5, 6, 8]

$$|\vec{n}, in\rangle = \sum_{\vec{s}} \delta_{|\vec{n}|, |\vec{s}|} \frac{\text{per}(U[\vec{n}|\vec{s}])}{\sqrt{\mu(\vec{n})\mu(\vec{s})}} |\vec{s}, out\rangle, \quad (2)$$

where  $\mu(\vec{n}) = \prod_{i=1}^M n_i!$ ,  $\text{per}(\dots)$  stands for the matrix permanent [4], and we denote by  $U[\vec{n}|\vec{s}]$  the  $N \times N$ -dimensional matrix obtained from the network matrix  $U$  by taking the  $k$ th row  $n_k$  times and the  $l$ th column  $s_l$  times (the order of rows/columns being unimportant). The probability of  $N_o$  clicks of the output detectors located at  $\vec{l} = (l_1, \dots, l_{N_o})$  reads

$$P_{out}(\vec{m}) = \sum_{\vec{s}} P_D(\vec{m}|\vec{s}) \sum_{\vec{n}} P_U(\vec{s}|\vec{n}) P_I(\vec{n}), \quad (3)$$

where the binary “occupation numbers”  $m_l$  count the detector clicks ( $m_{l_\alpha} = 1$  for  $1 \leq \alpha \leq N_o$  and  $m_{l_\alpha} = 0$  for  $N_o + 1 \leq \alpha \leq M$ ). Here the probability  $P_I(\vec{n})$  of the input  $\vec{n}$ , the conditional probability of the network output  $\vec{s}$ ,  $P_U(\vec{s}|\vec{n})$ , and the conditional detection probability  $P_D(\vec{m}|\vec{s})$  are given as follows:

$$P_U(\vec{s}|\vec{n}) = |\langle \vec{s}, out | \vec{n}, in \rangle|^2 = \frac{|\text{per}(U[\vec{n}|\vec{s}])|^2}{\mu(\vec{n})\mu(\vec{s})} \delta_{|\vec{n}|, |\vec{s}|},$$

$$P_I(\vec{n}) = \prod_{i=1}^N p_{n_i}, \quad P_D(\vec{m}|\vec{s}) = \prod_{l=1}^M P_D(m_l|s_l),$$

$$P_D(m|s) = e^{-\nu}r^s \delta_{m,0} + (1 - e^{-\nu}r^s) \delta_{m,1}. \quad (4)$$

We have obvious identities:  $\sum_{\vec{n}} P_I(\vec{n}) = 1$ ,  $\sum_{\vec{s}} P_U(\vec{s}|\vec{n}) = 1$ , and  $\sum_{\vec{m}} P_D(\vec{m}|\vec{s}) = 1$ .

Our model can also be extended to the recently proposed BS computer with Gaussian states [27], where a number,  $N_p$ , of the parametric down conversion sources is heralded for single photons, with a high probability that  $N$ , different at each run, input modes contain the state (1) in the modes  $i_1, \dots, i_N$  (in this case  $p_0 = 0$  and  $N = f(N_p)$ ). The difference with the original BS computer of Ref. [1] is that one samples both on the input and output modes.

We will focus on the network in the “collision free” limit  $M \gg N^2$  [1], when the BS computer is shown

to be classically hard to simulate. Due to the boson birthday paradox [1, 28, 29], the probability of photon bunching at the network output is bounded, on average in the Haar measure, by  $1 - (\sum_{|\vec{m}|=N} 1) / (\sum_{|\vec{n}|=N} 1) = 1 - \prod_{k=1}^{N-1} (1 - \frac{k}{M}) < \frac{N(N-1)}{2M}$  (for  $m_l \leq 1$ ). Thus the simple bucket detectors, registering only the presence of the input different from the vacuum, are sufficient. Similarly, it can be shown that the probability of a bunched output  $\vec{s}$  (i.e., some  $s_l > 1$ ),  $P_B(\vec{s}) = \sum_{|\vec{n}|=N_i} P_U(\vec{s}|\vec{n}) P_I(\vec{n})$ , in our case is bounded by  $\frac{N_i(N_i-1)}{2M}$ , on average in the Haar measure, where the overline denotes the averaging with respect to the probability of  $N_i$  photons in the input, i.e.,  $P_I(N_i) \equiv \sum_{|\vec{n}|=N_i} P_I(\vec{n})$ .

*The variational distance to the ideal BS computer* – The variational distance between the output distributions of the ideal and an imperfect BS devices is the measure of the computational complexity of the latter. For  $M \gg N^2$  we can ignore the bunched output in the ideal BS case, thus the variational distance to the ideal BS computer  $\mathcal{V}$  consists of the following two parts:

$$\mathcal{V}_1 \equiv \sum_{|\vec{m}| \neq N} P_{out}(\vec{m}), \quad \mathcal{V}_2 \equiv \sum_{|\vec{m}|=N} |P_{out}(\vec{m}) - P_{out}^{(0)}(\vec{m})|, \quad (5)$$

where  $P_{out}^{(0)}(\vec{m}) \equiv P_U(\vec{m}|\vec{n}^{(0)})$  and we have introduced the ideal input  $\vec{n}^{(0)}$ , i.e.,  $n_i^{(0)} = 1$  for  $i = 1, \dots, N$  and  $n_i^{(0)} = 0$  otherwise. First of all, in  $\mathcal{V}_{1,2}$  of Eq. (5) we have an exponential number of terms in the summation over  $\vec{m}$  and over  $\vec{s}$  in  $P_{out}(\vec{m})$  of Eq. (3) (and also over  $\vec{n}$ ). An exponentially small bound on the probability  $P_U(\vec{s}|\vec{n})$  is needed to bound such a sum. Whereas there seem to be no deterministic condition on the  $U$ -matrix elements to have all output probabilities  $P_U(\vec{s}|\vec{n})$  exponentially bounded as needed for our purposes, there is a probabilistic bound in the Haar measure: we simply exclude a fraction  $\delta$  of the networks. One way of doing it is to use Chebyshev's inequality which bounds the tail of a probability distribution by its moment.

The easiest to compute moment of the variational distance  $\mathcal{V}$  (5), for the Haar-random  $U$ , is its average value. For  $M \gg N^2$  any  $N \times N$ -dimensional submatrix of a Haar-random  $M \times M$ -dimensional  $U$  is made of the elements approximated by the i.i.d. complex Gaussian random variables with the probability density  $p(U_{kl}) = \frac{M}{\pi} \exp\{-M|U_{kl}|^2\}$  [1]. A simple way to obtain the average of the probability  $P_U(\vec{s}|\vec{n})$  for arbitrary  $\vec{s}$  and  $\vec{n}$  is to use a formula for the matrix permanent employing the Fisher-Yates distribution of the contingency tables  $T$ :  $\mathcal{P}(T|\vec{s}, \vec{n}) = \frac{\mu(\vec{s})\mu(\vec{n})}{N_i! \mu(T)}$ , where  $N_i = |\vec{s}| = |\vec{n}|$ ,  $T$  is a  $M \times M$ -dimensional matrix such that  $\sum_{l=1}^M T_{kl} = n_k$  and  $\sum_{k=1}^M T_{kl} = s_l$  (the contingency table), and  $\mu(T) = \prod_{k,l=1}^M T_{kl}!$ . From Ref. [30] we have

$$\text{per}(U[\vec{s}|\vec{n}]) = N_i! \sum_T \mathcal{P}(T|\vec{s}, \vec{n}) \prod_{k,l=1}^M U_{kl}^{T_{kl}}. \quad (6)$$

Using the Gaussian approximation we get  $\langle \prod_{k,l=1}^M U_{kl}^{T_{kl}} (U_{kl}^{T_{kl}})^* \rangle = \delta_{T,T'} \frac{\mu(T)}{M^{N_i}}$  (where  $\langle \dots \rangle$  stands for the average over  $U$ ). From Eqs. (4) and (6) we get

$$\langle P_U(\vec{s}|\vec{n}) \rangle = \frac{N_i!}{M^{N_i}} \delta_{|\vec{s}|, N_i} \delta_{|\vec{n}|, N_i}, \quad (7)$$

valid for  $N_i^2 \ll M$ . In the case of small errors, the average number of photons  $N_i$  in the input is close to  $N$ , hence, the Gaussian approximation can be still used in the calculations below [33]. By using Eq. (7) for the averaging over  $U$  we obtain

$$\begin{aligned} \langle \mathcal{V}_1 \rangle &= 1 - \sum_{|\vec{m}|=N} \sum_{\vec{s}} P_D(\vec{m}|\vec{s}) \sum_{\vec{n}} \langle P_U(\vec{s}|\vec{n}) \rangle P_I(\vec{n}) \\ &= 1 - \sum_{|\vec{m}|=N} \sum_{N_i=0}^{\infty} \sum_{|\vec{s}|=N_i} P_D(\vec{m}|\vec{s}) \frac{N_i!}{M^{N_i}} \sum_{|\vec{n}|=N_i} P_I(\vec{n}) \\ &\leq 1 - e^{-(M-N)\nu} (1 - e^{-\nu} r)^N p_1^N \left[ 1 - \frac{N^2}{2M} \right] \\ &\equiv 1 - Q \left[ 1 - \frac{N^2}{2M} \right]. \end{aligned} \quad (8)$$

We have retained only the terms with  $N_i = N$  from the sum over  $\vec{s}$  in Eq. (8), used that  $\sum_{|\vec{n}|=N} P_I(\vec{n}) \geq p_1^N$ , and the following inequality

$$\begin{aligned} &\sum_{|\vec{m}|=N} \sum_{|\vec{s}|=N} P_D(\vec{m}|\vec{s}) \frac{N!}{M^N} \\ &\geq \frac{M!}{M^N (M-N)!} e^{-(M-N)\nu} (1 - e^{-\nu} r)^N \\ &\geq \left[ 1 - \frac{N^2}{2M} \right] e^{-(M-N)\nu} (1 - e^{-\nu} r)^N, \end{aligned} \quad (9)$$

where the single term  $P_D(\vec{m}|\vec{m})$  is retained to bound from below the sum over  $\vec{s}$  (which is also reasonably close to the whole sum for small errors), taken into account that the number of all outputs  $\vec{m}$  is  $\frac{M!}{N!(M-N)!}$ , and the fact that  $\frac{M!}{(M-N)!} > M^N \left[ 1 - \frac{N^2}{2M} \right]$ . On the r.h.s. of Eq. (8) we subtract from 1 the product of the bound  $1 - \frac{N^2}{2M}$  on the average probability of the non-bunched output and  $Q$  – the probability that  $N$  detectors have clicked,  $M - N$  detectors had zero dark counts, and that  $N$  indistinguishable single photons are at the network input.

We split  $\mathcal{V}_2$  into three parts according to the terms in the sums over  $\vec{s}$  and  $\vec{n}$  in Eq. (3). By abusing the notations slightly, we have

$$\mathcal{V}_2 \leq \mathcal{V}_2 \left[ \begin{array}{c} \vec{s} = \vec{m} \\ \vec{n} = \vec{n}^{(0)} \end{array} \right] + \mathcal{V}_2 \left[ \begin{array}{c} \vec{s} \neq \vec{m} \\ \vec{n} = \vec{n}^{(0)} \end{array} \right] + \mathcal{V}_2 \left[ \begin{array}{c} \forall \vec{s} \\ \vec{n} \neq \vec{n}^{(0)} \end{array} \right], \quad (10)$$

where the first term on the r.h.s. of Eq. (10) contains  $P_{out}^{(0)}(\vec{m})$ . By using Eq. (7) we get (noticing that the first

term on the r.h.s., due to  $P_{out}^{(0)}(\vec{m})$ , is larger)

$$\begin{aligned} \left\langle \mathcal{V}_2 \left[ \begin{array}{c} \vec{s} = \vec{m} \\ \vec{n} = \vec{n}^{(0)} \end{array} \right] \right\rangle &= \sum_{|\vec{m}|=N} \left\{ \left\langle P_U(\vec{m} | \vec{n}^{(0)}) \right\rangle - \right. \\ &\quad \left. - P_D(\vec{m} | \vec{m}) \left\langle P_U(\vec{m} | \vec{n}^{(0)}) \right\rangle p_1^N \right\} \\ &= \sum_{|\vec{m}|=N} [1 - P_D(\vec{m} | \vec{m}) p_1^N] \frac{N!}{M^N} \leq 1 - Q, \end{aligned} \quad (11)$$

where we have identified  $Q$  of Eq. (8). Similarly as in Eqs. (8)-(9), we obtain

$$\begin{aligned} \left\langle \mathcal{V}_2 \left[ \begin{array}{c} \vec{s} \neq \vec{m} \\ \vec{n} = \vec{n}^{(0)} \end{array} \right] \right\rangle &= 1 - \sum_{|\vec{m}|=N} \left\{ P_D(\vec{m} | \vec{m}) p_1^N \right. \\ &\quad \left. \times \left\langle P_U(\vec{m} | \vec{n}^{(0)}) \right\rangle \right\} \leq 1 - Q \left[ 1 - \frac{N^2}{2M} \right]. \end{aligned} \quad (12)$$

Finally, using the identities  $\sum_{\vec{m}} P_D(\vec{m} | \vec{s}) = 1$  and  $\sum_{\vec{s}} P_U(\vec{s} | \vec{n}) = 1$ , we obtain for the last term in Eq. (10)

$$\begin{aligned} \mathcal{V}_2 \left[ \begin{array}{c} \forall \vec{s} \\ \vec{n} \neq \vec{n}^{(0)} \end{array} \right] &= \sum_{|\vec{m}|=N} \sum_{\vec{s}} \sum_{\vec{n} \neq \vec{n}^{(0)}} \left\{ P_D(\vec{m} | \vec{s}) P_U(\vec{s} | \vec{n}) \right. \\ &\quad \left. \times P_I(\vec{n}) \right\} \leq \sum_{\vec{n} \neq \vec{n}^{(0)}} P_I(\vec{n}) = 1 - p_1^N \equiv Q'. \end{aligned} \quad (13)$$

Gathering together the contributions (8) and (11)-(13) we obtain an upper bound on the Haar-average variational distance (5) (valid for  $M \gg N^2$ )

$$\langle \mathcal{V} \rangle \leq 2 \left\{ 1 - Q \left[ 1 - \frac{N^2}{2M} \right] \right\} + 1 - Q + Q' \equiv \mathcal{R}_A. \quad (14)$$

By employing Chebyshev's inequality for the Haar probability measure  $Pr(\dots)$ , which in this case reads  $Pr(\mathcal{V} < \epsilon) \geq 1 - \langle \mathcal{V} \rangle / \epsilon$ , one can deduce a sufficient condition that the experimental BS device is  $\epsilon$ -close in the variational distance to the ideal BS computer for the fraction  $1 - \delta$  of the network matrices:

$$\mathcal{R}_A(r, \nu, p_1) \leq \epsilon \delta. \quad (15)$$

For small imperfections  $M\nu \ll 1$ ,  $Nr \ll 1$ , and  $N(1 - p_1) \ll 1$  (and  $M \gg N^2$ ) we have  $\mathcal{R}_A \approx 2 \left( 1 - \frac{N^2}{2M} \right) + 3[(M - N)\nu + Nr] + 4N(1 - p_1)$ . Hence, a sufficient condition for the bound in Eq. (15) reads

$$1 - \frac{N^2}{2M} + \frac{3}{2} [(M - N)\nu + Nr] + 2N(1 - p_1) \leq \frac{\epsilon \delta}{2}. \quad (16)$$

Eq. (16) tells us that if the experimental parameters  $\nu$ ,  $r$ , and  $p_1$  satisfy  $\nu = O(M^{-1})$ ,  $r = O(N^{-1})$ , and  $p_1 = 1 - O(N^{-1})$  then the corresponding experimental BS device is scalable, i.e., for any  $N$  it has a constant variational distance error  $\epsilon$  to the ideal BS computer, with the success probability  $1 - \delta$  in the Haar measure.

The above derived scalability conditions can be supplemented to the complete set by using the results of Refs. [22, 23]. In Ref. [22] it was shown that the fidelity of optical elements in the network must be at least  $1 - O(N^{-2})$  for the noisy-network BS device to be scalable (one can also check the network unitarity *in situ*, see below). Using the same notations for the variational distance error  $\epsilon$  and the success probability  $1 - \delta$ , the following scalability condition on the photon mode mismatch is sufficient (see Eq. (26) in Ref. [23]) [34]

$$\mathcal{R}_B(\vec{g}) \leq \epsilon^2 \delta, \quad (17)$$

where the indistinguishability parameters  $\vec{g} = (g_2, \dots, g_N)$  of the single photons are defined by the density matrix of the photon source projected onto the one-particle subspace ( $\rho_1$  in the above notations). We have [23]  $g_k \equiv \text{Tr}(\rho_1^k) / \text{Tr} \rho_1$  and

$$\mathcal{R}_B(\vec{g}) \equiv \sum_{\vec{c}} \frac{\chi(c_1) \left( 1 - \prod_{k=2}^N g_k^{c_k} \right)^2}{\prod_{k=1}^N k^{c_k} c_k!}, \quad (18)$$

where the summation is over  $\vec{c} = (c_1, \dots, c_N)$  satisfying  $\sum_{k=1}^N k c_k = N$  and  $\chi(n) = \sum_{k=0}^n \frac{n!}{k!} = \int_1^\infty dz z^n e^{1-z}$ . For small mode mismatch [23]

$$\mathcal{R}_B \approx (1 - \langle \mathcal{F}_{ph} \rangle)^2 \left( \frac{N^3}{3} - \frac{N^2}{2} + \frac{7N}{6} - 1 \right) \leq \epsilon^2 \delta \quad (19)$$

leading to the scalability condition  $1 - \langle \mathcal{F}_{ph} \rangle = O(N^{-3/2})$ , where the average fidelity of the single photons reads  $\langle \mathcal{F}_{ph} \rangle \approx (1 + g_2)/2$ .

*Verification of the experimental BS device* – We have shown that, for a unitary network of size  $M$  with  $N$  photon sources at the input (in the dilute limit  $M \gg N^2$ ), given a variational distance error  $\epsilon$ , such an experimental BS device outputs a classically hard probability distribution with the success probability  $1 - \delta$ , if the conditions (15)-(16) and (17)-(19) are satisfied. Note that to ensure a classically hard instance of  $U$  one simply has to use  $1/(1 - \delta)$  randomly chosen networks, on average.

Assuming that an operational device is available, i.e., there is a black box claimed to be the BS computer, how one could verify it? Generally, to certify *unconditionally* (i.e. without any additional assumptions not verified in the test itself) that an experimental device simulates the BS computer output, the test must be non-polynomial for the classical computing. Otherwise, there is a black-box simulator (e.g., a program on a computer) which would pass the test. In this respect, the variational distance, which quantifies the complexity of an experimental BS device by its closeness to the ideal BS computer, can also serve as such an unconditional test. The variational distance error can be obtained by comparison with the classical simulations of the ideal BS computer, feasible for up to  $N \sim 30$  photon sources.

The unconditional variational distance test requires an exponential number of the experimental runs. That is

why the *conditional* tests, i.e., the tests based on additional conditions, are important as some evidence of the BS computer operation. Such tests can be conditioned on some features of the experimental setup. For instance, the unitarity of the network can be checked *in situ*. Indeed, the linear map  $\varphi(U)$  defined in Eq. (2) is also unitary [8] and preserves the group property:  $\varphi(U_2 U_1) = \varphi(U_2) \varphi(U_1)$ . Therefore, for any input  $|\Psi, in\rangle$ , the map  $|\Psi, out\rangle = \varphi(U)|\Psi, in\rangle$  is invertible, with the inverse map given by  $\varphi(U^\dagger) = [\varphi(U)]^\dagger$ . Hence, by placing the high-quality mirrors at the output of the network (instead of the detectors) one can redirect the photons to pass through the  $U^\dagger$ -network after the  $U$ -network, resulting in their return to the same input modes. One simply has to check the absence of the photons in the inputs  $N+1, \dots, M$  by using the bucket detectors. This test verifies if the first-order coherence is still preserved in the output of the network. Furthermore, if the input is certified to satisfy the above derived scalability conditions (i.e., the input is close to the indistinguishable

single photons) the unitarity test for a random network is a conditional test of the BS device operation, since the photons must pass the output state of the  $U$ -network having a classically hard probability distribution. One can devise other, more sophisticated, conditional tests of the BS device operation for some specific networks with symmetries. For instance, by using the  $N$ -th order generalization of the HOM effect [13] one can check that the  $N$ -th order coherence is preserved, as is recently proposed in Ref. [31]. This test is also a conditional test, since it verifies the needed  $N$ -th order coherence *in situ* but, on the other hand, it is only polynomial in  $N$  because it simply checks for the zero probability in some of the output configurations and is independent of the distribution in all other output configurations.

This work was supported by the CNPq of Brazil. Helpful discussions with M. C. Tichy in the initial stage of this work are acknowledged. The author is also grateful to P. P. Rohde for a valuable comment.

- 
- [1] S. Aaronson and A. Arkhipov, arXiv:1011.3245 [quant-ph]; Theory of Computing **9**, 143 (2013).
  - [2] P. Shor, in *Proceedings of the 35th Annual Symposium on Foundations of Computer Science* (IEEE Comp. Soc. Press, Los Alamos, CA, 1994), p. 124; SIAM J. Comput. **26**, 1484 (1997).
  - [3] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge Univ. Press, 2000).
  - [4] H. Minc, *Permanents, Encyclopedia of Mathematics and Its Applications*, Vol. **6** (Addison-Wesley Publ. Co., Reading, Mass., 1978).
  - [5] E. R. Caianiello, Nuovo Cimento, **10**, 1634 (1953); *Combinatorics and Renormalization in Quantum Field Theory*, Frontiers in Physics, Lecture Note Series (W. A. Benjamin, Reading, MA, 1973).
  - [6] S. Scheel, arXiv:quant-ph/0406127.
  - [7] L. G. Valiant, Theoretical Comput. Sci., **8**, 189 (1979).
  - [8] S. Aaronson, Proc. Roy. Soc. London A, **467**, (2008) 3393.
  - [9] H. Ryser, *Combinatorial Mathematics*, Carus Mathematical Monograph No. 14. (Wiley, 1963).
  - [10] See, for instance, the reviews: B. Lounis and M. Orrit, Rep. Prog. Phys. **68**, 1129 (2005); G. S. Buller and R. J. Collins, Meas. Sci. Technol. **21**, 012002 (2010); M. D. Eisaman, J. Fan, A. Migdall, and S. V. Polyakov, Rev. Sci. Instr. **82**, 071101 (2011).
  - [11] C. K. Hong, Z. Y. Ou, and L. Mandel, Phys. Rev. Lett. **59** (1987) 2044.
  - [12] Y. L. Lim and A. Beige, New J. Phys., **7** 155 (2005).
  - [13] M. C. Tichy *et al*, New J. Phys., **14** (2012) 093015.
  - [14] M. A. Broome *et al*, Science **339**, 794 (2013).
  - [15] J. B. Spring *et al*, Science, **339**, 798 (2013).
  - [16] M. Tillmann *et al*, Nature Photonics, **7**, 540 (2013).
  - [17] A. Crespi *et al*, Nature Photonics, **7**, 545 (2013).
  - [18] D. Aharonov and M. Ben-Or, In Proc. ACM STOC p. 176 (1997); quant-ph/9703054.
  - [19] E. Knill, R. Laflamme, and W. Zurek, Science **279**, 342 (1998).
  - [20] P. P. Rohde and T. C. Ralph, Phys. Rev. A **85**, 022332 (2012).
  - [21] P. P. Rohde, Phys. Rev. A **86**, 052321 (2012).
  - [22] A. Leverrier and R. García-Patrón, arXiv:1309.4687 [quant-ph].
  - [23] V. S. Shchesnovich, Phys. Rev. A **89**, 022333 (2014).
  - [24] P. P. Rohde, K. R. Motes, P. Knott, and W. J. Munro, arXiv:1401.2199 [quant-ph].
  - [25] S. M. Barnett, L. S. Phillips, and D. T. Pegg, Opt. Commun. **158**, 45 (1998).
  - [26] H. Lee *et al*, J. Mod. Opt. **51**, 1517 (2004).
  - [27] A. P. Lund *et al*, arXiv:1305.4346 [quant-ph].
  - [28] A. Arkhipov and G. Kuperberg, Geom. Topol. Monogr., **18**, 1 (2012).
  - [29] N. Spagnolo *et al*, Phys. Rev. Lett. **111**, 130503 (2013).
  - [30] V. S. Shchesnovich, Int. J. of Quant. Inf. **11**, 1350045 (2013).
  - [31] M. C. Tichy, K. Mayer, A. Buchleitner, and K. Molmer, arXiv:1312.3080 [quant-ph].
  - [32] The additional parameter  $\delta$  is the fraction of the network matrices in the Haar measure to which the conditions below are not applicable.
  - [33] Since  $N_i$  can be on the order or even larger than  $\sqrt{M}$ , there is a success probability  $1 - \delta_i$  when Eq. (7) is used for unbounded  $N_i$ . We assume  $\delta_i < \delta$ , where the success probability  $1 - \delta$  is introduced below.
  - [34] There is a change of notations: the variational distance is scaled by 2 and our  $\mathcal{R}_B(\bar{g})$  is equal to the  $\mathcal{V}(N, \eta)$  of Ref. [23].